XZY Authentication Agency

Under the EU AI Act, the concept of "substantial human editorial control" serves as a critical exemption. If an organization or creator uses AI to generate text on matters of public interest, they are generally required to disclose that the text is AI-generated. However, Article 50(4) explicitly waives this disclosure requirement if the content has undergone a process of human review and a designated person assumes editorial responsibility for it.

The law and the recently finalized European Commission Guidelines make it very clear: this exemption is not a loophole for rubber-stamping AI output. The human oversight must be genuine, meaning the human must have the authority and the actual

opportunity to alter, correct, or reject the AI's work before publication.

Here is how the EU AI Act defines this control and the documentation required to prove it.

Defining "Genuine" Editorial Control

To qualify for the editorial control exception, the process must meet two criteria:

1. **Substantive Review: The human review cannot be limited to superficial matters, routine checks, or cursory pro-forma approval. The reviewer must engage with the semantic meaning, factual accuracy, and context of the AI-generated text.**
2. **Accountability: A natural or legal person must formally hold editorial responsibility for the publication of the content. This means someone's name or a specific corporate entity's reputation is legally tied to the output, assuming the same liability as if a human had drafted the piece entirely from scratch.**

Essentially, the human must act as an active co-creator or rigorous editor, not merely a passive conduit for the AI's output.

Documentation and Proof of Oversight

While the AI Act does not prescribe a rigid, standardized form that must be filled out for every single paragraph generated, organizations must be able to demonstrate to market surveillance authorities that their editorial control is real and systematic. If an undisclosed AI

article is flagged by a regulator, the publisher must be able to prove the human-in-the-loop process.

To establish an auditable record of human oversight, organizations typically need to maintain the following documentation:

- **Workflow Policies: Written internal policies detailing the mandatory steps a piece of AI-assisted content must go through before publication. This should define who the responsible editors are and what criteria they use to evaluate the content.**
- **Version History/Audit Trails: System logs or document version histories (like tracked changes in a word processor or a CMS revision history) that show the progression from the raw AI output to the final published version. This proves that a human actually modified, corrected, or actively validated the text.**
- **Sign-off Records: A clear, recorded approval (e.g., in a content management system or project management tool) linking a specific, identifiable human editor to the final approval of the piece.**

  Important Distinction: This editorial exception applies only to AI-generated text. It does not apply to audio, video, or image generation (deepfakes or synthetic media). If you generate a synthetic video or audio file, it must be labeled regardless of how much human review went into approving it.

ARE U AI PROOF?